

Design and evaluation of an interactive proof-of-concept dashboard for general practitioners

Robin De Croon, Joris Klerkx, Erik Duval
KU Leuven
Celestijnenlaan 200A
BE-3001 Heverlee

Email: robin.decroon@cs.kuleuven.be, joris.klerkx@cs.kuleuven.be, erik.duval@cs.kuleuven.be

Abstract—Targeted follow-up meetings in general practice are important and missed often, because of both patient and general practitioners (GPs) related reasons. In this paper, we present a proof-of-concept interactive visualization dashboard that provides GPs with a powerful, yet easy to use method to identify those patients in need of follow-up. We applied a user centered, rapid prototyping methodology with 12 information visualization students and 15 GPs. We evaluated the final design using the evaluation framework by O’Leary *et al.* [1], as well as a System Usability Scale questionnaire. Results indicate that there is indeed a need for a follow-up tool and that a dashboard is a right kind of tool. Our proof-of-concept shows useful insights into patient records and can indeed help GPs recognize patients in need of follow-up. The major strengths of the design are the ease with which GPs can query patient records using interactive visualizations, such as parallel coordinates, and the ability to check if the number of patients diagnosed with certain diseases differs from the amount predicted in evidence-based guidelines.

I. INTRODUCTION

According to guidelines of many health conditions, such as for instance hypertension and high cholesterol, patients are required to re-visit a general practitioner (GP) for a follow-up meeting after diagnosis.¹ In a follow-up meeting, a physician typically checks the health status and registers any changes or events that took place after the patient’s last visit. In health domains such as dentistry and physiotherapy, follow-up meetings are common practice. However, in general practice, patients often miss incentive and are sometimes asked to take responsibility and arrange a meeting themselves [2]. But, patients often forget this, since they do not consider a follow-up necessary. This can be explained with patient oriented problems: 1) patients do not want to spend additional money; 2) patients simply forget or get confused; 3) patients have the impression the treatment plan is not working; or by GP oriented problems: 4) GPs are too busy to accommodate a prompt visit; and 5) not enough time is spend with the GP.^{2 3} Although these problems should be addressed, an appropriate system to detect all patients in need of a follow-up appointment is also needed. In some electronic medical record (EMR) systems, notifications are shown when patients check in for another visit. Nevertheless, this is inadequate, since only patients are addressed who attend for another

consultation. Missed follow-up meetings result in missed or delayed diagnoses [2].

In contrast, as a visit consumes both time and money, too many follow-ups are also unwanted. Ganguli *et al.* [3] stress the importance of targeted follow-up meetings, saving money and resources on unnecessary appointments. So, two possible methods arise in order to improve patient follow-up quality: 1) a patient-oriented approach in which patient empowerment can be stimulated; or 2) a GP driven approach in which GPs are provided with better tools to detect patients in need of follow-up consultations. Of course, the ideal solution is to improve patient follow-up combining both methods. Nevertheless, this paper specifically focuses on the second approach by providing GPs with an interactive visualization dashboard that helps them detect relevant patients. We do not want to replace GPs, yet we want to augment their tool set, which is in line with important previous research as described by Engelbart [4]: “By ‘augmenting human intellect’ we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems.”.

Some electronic medical record (EMR) systems offer rather basic options in order to gain insights in patient follow-up. For instance, GPs can typically only filter patient records based on simple parameters such as length, weight, blood pressure and BMI. However, these characteristics do not directly relate to health conditions that require follow-up meetings. Filtering patient records based on these parameters requires the GP to run pre-configured search queries that do not allow for fine-grained control. Furthermore, GPs first need to manually run the query before they know whether there are patients within the requested range, whereas a visual solution shows this in advance. Moreover, with a visual solution, GPs do not need to know what they are looking for in advance. For instance, when GPs wants to check all patients with a BMI higher than 30, they would most likely not consider that there are some patients with a BMI lower than 15. Traditional tools do not notify GPs of values they are not looking for, while in a well-designed visual solution the user can immediately see these values exist.

Because of their limited options, GPs who want to know about their follow-up quality in general request an audit from the Academic Center of General Practice. However, not many GPs request an audit as this is often perceived as a control system. Therefore, the work presented in this paper tries to empower GPs with interactive visualizations in which they can

¹nlm.nih.gov/medlineplus/ency/article/000468.htm

²<http://medicaloffice.about.com/od/patientsatisfaction/tp/>

³Reasons-Patients-Dont-Come-Back.htm

³<http://well.blogs.nytimes.com/2014/11/13/when-patients-dont-follow-up/>

use visual filters to reflect on their patient records. Since a large practice easily consists of several thousand patients, it is impossible to contact all patients for all possible conditions. In order for GPs to stay in control and enable them to still focus on a broad range of patients, we present a thoroughly designed visual query tool that allows GPs to easily filter out patient records of interest.

The remainder of the paper is structured as follows. Section II situates the presented work in literature and outlines the benefits of a visual approach. Section III describes our iterative design process, whereas Section IV concisely presents some implementation details. The final design is evaluated and discussed in Section V and Section VI respectively. Finally the paper concludes with conclusions and future work in Section VII.

II. LITERATURE REVIEW

This section starts with a short overview of related work. We are not aware of other visual tools that focus specifically on patient follow-up meetings in general practice. Below, we provide an overview of the most relevant medical information visualization tools discussed in literature.

Caban and Gotz [5] have stated in a recent (2015) editorial of the Journal of the American Medical Informatics Association that “[t]he large volume of clinical data now being captured for each patient poses many challenges to clinicians trying to combine data from different disparate systems and make sense of the patient’s condition within the context of the patient’s medical history” [5]. A possible approach to deal with these large amounts of data are data visualizations; several visual analytics tools and criteria dealing with medical data are extensively discussed by Rind *et al.* [6]. Even more recently West *et al.* [7] presented a systematic review paper that groups over 800 articles that contain some kind of data visualization in the domain of healthcare informatics. Their research lists ten challenges that needs to be addressed [7]. In our research we aim to address five of these challenges: 1) better options to manage large amounts of data are needed; 2) electronic health record data are complex and large in size; 3) visualizing data is important to users; 4) medical data is even more complicated due to missing values and inaccurate entries; 5) presenting information interactively in a single screen size is an important design feature.

Ben Shneiderman stressed the importance of visualizations in healthcare during his opening speech at EHRVis 2014,⁴ while Dasgupta *et al.* [8] stressed the importance of privacy. Since our dashboard only visualizes data that GPs already have access to, this paper will not go into detail on privacy issues. In addition, Malik *et al.* [9] presented CoCo, which is a visual analytics tool with integrated support for statistical analysis. Some results from CoCo are similar outcomes to the work presented in this paper: 1) improve collaboration among colleagues; 2) improve intermediate result discussion; and 3) create meaningful outcome presentation [10]. Nevertheless, the goal is different as we primarily aim to help GPs gain insight into their patient follow-up.

LifeLines [11] visualizes patient summaries using graphical attributes, such as colors and lines depicting a patient’s discrete

events. LifeLines2 [12] and EventFlow [13] displays and summarizes time-point and interval data; individuals with the same sequence of data points are grouped and the average interval time between these events is visualized [13]. Similar to LifeLines, VISITORS [14], or Visualization of Time-Oriented Records, accommodates diverse temporal data from multiple records. A more algorithmic approach is developed by Gotz *et al.* [15], called Dynamic Icons, or DICON. It is designed to interactively explore clusters of similar patients in a treemap. Once users understood the principles of a treemap, the interface could help them to analyze data more rapidly. We hope that, similar to DICON, once users are familiar with the visualizations used in our research, they help GPs analyze their patient records as well.

Ratwani and Fong [16] present several challenges such as overplotting and the uncertainty how to discriminate and visualize all different factors over time to expose structures of interest in the data. They developed a visual mining system based on OutFlow [17] to support exploratory analysis of electronic medical records. The OutFlow [17] visualization technique is basically a sankey diagram [18], which is designed to visualize flow processes. Sankey diagrams, however, primarily focus on the proportion that splits in different paths without temporal information [19]. Furthermore, the visual display of Outflow looks similar to parallel coordinates [20], which we use in our design too. However, the underlying data types are different. Outflow visualizes temporal event sequences and our parallel coordinates visualize quantitative data, as discussed in more detail in Section III. Ratwani and Fong’s [16] work is similar to ours in the sense that they also allow users to quickly filter, and interact with the data. However, we focus on general practice instead of a hospital setting. Another study on patient follow-up in a hospital setting is done by Jodeiry *et al.* [21], yet the major difference between these two contexts is the management of patients. In general practice GPs typically manage their own patient records based on a more personal relationship whereas in a hospital patients are managed using a central system. A more thorough review of other dashboards in a clinical settings for improving patient care can be found in [22].

Contrary to the multi-patient approach applied in our research, a good single patient visualization is designed by Zhang *et al.* [23], called AnamneVis. It is based on the Five Ws concept for information gathering: who, when, what, where, and why. The patient (who) is visualized using a radial sunburst visualization. The other Ws are part of a reasoning chain that can be interactively sorted and brushed [23]. They conclude that a visual solutions such as AnamneVis can “*significantly lower the time and effort needed to access the medical patient information required to arrive at a diagnostic conclusion.*”. We also aim to lower the burden to access information stored in medical patient records, but for multiple patient records simultaneously.

III. ITERATIVE DESIGN

In our study we applied a user-centered, rapid-prototyping methodology, similar to our previous work [24] and to other researchers such as Ratwani and Fong [16]. In line with this methodology, our prototype evolved quickly. After every evaluation with real end users (i.e. GPs), problems

⁴<http://www.cs.umd.edu/hcil/parisehrvis/>

were addressed in the next prototype, which was then again evaluated. This section starts with explaining the rationale of the initial design. Secondly, the perceived usefulness and usability of this initial design is evaluated within the Academic Center for General Practice. Thirdly, the updated design is evaluated with GPs not affiliated with any academic center and the design is thereafter evaluated for usability issues during an ‘information visualization’ course. Lastly, all remarks are integrated into the final design.

A. Initial design

Our design follows Shneiderman’s information seeking mantra [25]: “*overview first, zoom and filter, then details on demand*”. As shown in Fig. 1, our dashboard consists of several widgets that serve as visual filters or dynamic queries [26]. As such, all patients are initially visualized and GPs can manually filter patients based on their criteria. Patients who require follow-up should be identified quickly in the overview. Moreover, no predetermined filters are set since we want to start with the overview first. Finally, since GPs have limited free time available, the design aims to assist GPs as much as possible in detecting interesting ranges and patterns in their data rapidly themselves.

Following Stevens’ model [27], EMR data can be divided into different data types: nominal data, such as information on the main caregiver or medication, and quantitative measurements or parameters, such as blood pressure or weight. These can be measured by GPs or in a laboratory. In order to efficiently filter patients based on these multivariate parameters, a parallel coordinates widget is added, as can be seen in part A of Fig. 1. Parallel coordinates are a good visualization technique to visualize and analyze multivariate data [20]. Every parameter corresponds to vertical, uniformly spaced axes, and a patient is visualized as a connected set of points, one for each parameter on every axis. The initial parameters are similar to the basic set of parameters present in existing EMR systems: systolic and diastolic blood pressure, weight, blood sugar, and age. This basic set is extended with income and ‘days since last visit’ in order to discover correlations between income [28] and follow-up. Parallel coordinates are a space-efficient method to represent large multi-variate data-sets. The state of the art of the parallel coordinates visualization technique is described by Heinrich and Weiskopf [29]. They cover models for constructing parallel coordinates, review methods for creating and understanding visual representations of parallel coordinates, and discuss challenges in constructing and understanding parallel coordinates. Furthermore, interactivity can be added easily with brushing and linking techniques: the desired ranges can be selected by brushing over the respective axis as shown in Fig. 2. Brushing can be applied on multiple parameters simultaneously. For example, GPs can select patients with hypertension, which are younger than 50 years and have a heart-rate higher than 100 bpm. The parallel coordinates show how selected patients perform on other parameters as well. Furthermore, thank to the linked views, for every visual filter that is applied by selecting a certain range on an axis, all other widgets immediately update to the current selection.

The National Center for Health Statistics in the USA has

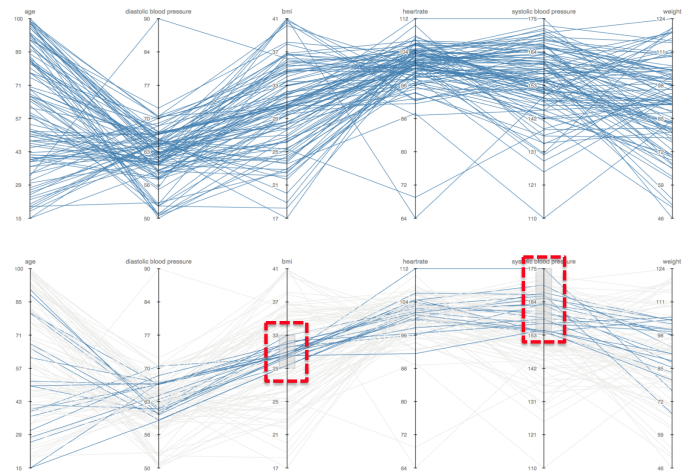


Fig. 2. Example on how GPs can brush parameters in parallel coordinates. In the top figure, no patients are selected. However, in bottom visualization all patients with a BMI between 29 and 33, and who have a systolic blood pressure higher than 153 mmHg are selected.

shown that population groups with the highest drug poisoning death rates in 2008 were males, aged 45–54 years, and non-Hispanic white and American Indians or Alaska Natives.⁵ Since these are typical characteristics GPs can use to find and predict problem areas, part B in Fig. 1 contains demographic widgets. The upper widget allows to easily filter on gender and age, and the map widget adds geographical information using a heatmap visualization. Using the map, GPs can check if there are problems in specific areas. For instance, geographical correlations can show if and when there is a flu outbreak in a certain region [30]. This enables GPs to identify and contact patients from specific areas as a preventive precaution. Moreover, geolocation is used in an on-going study at the Academic Center of General Practice of Leuven to show the prevalence of a chronic obstructive pulmonary disease in proximity of a steel manufacturer. Some GPs read such studies in order to keep up with new findings in medical research, therefore they can easily select all patients for follow-up from a particular area. The common bar charts are formatted according to Tufte’s design guidelines [31], such as optimizing the Data-Ink ratio, in order to allow GPs to filter on qualitative data, such as population groups.

Finally, the bottom part of the dashboard lists the filtered patients. Physical follow-up is not always needed; even a quick phone call is associated with increased patient satisfaction and resolution of medication-related problems [32]. Therefore, it should be possible to open the patient file, call the patient, send (custom) appointment reminders or flag a warning when the patient comes for another visit. Ganguli *et al.* [3] even mention that: “[a]utomated email follow-up and other types of communication could potentially replace a significant fraction of in-person encounters.”. This functionality is only conceptually explained since no actual link with the GP’s EMR system was made for our proof-of-concept.

⁵<http://www.cdc.gov/nchs/data/databriefs/db81.htm>



Fig. 1. Overview of the first prototype which exists only in Dutch. Initially data of all patients is visualized in the dashboard (overview first). Part A shows the parallel coordinates widget where each line represents a single patient. The GP can adjust filters by selecting an area on the representative axes. Part B shows demographic information of the filtered patients. In clockwise direction are shown: a population pyramid, most occurring conditions, religion, main caregiver, highest degree and a heatmap. The bar charts serve as ordinal filters. Part C lists the patients filtered by the visual queries (details on demand). The table shows from left to right: picture, name, condition, gender, age, degree and main GP. Note that all demo pictures are copyrighted under a CC BY-NC-SA 2.0 license.

B. Academic Center for General Practice iteration

The perceived usefulness and usability of the initial design has been checked with three GPs and two EMR experts at the Academic Center for General Practice in Leuven, which performs audits on general practices in Belgium to assess follow-up quality.

The proof-of-concept suffered from an entry barrier as GPs are not familiar with parallel coordinates. Yet, once the researcher shortly explained the principle of parallel coordinates, the proof-of-concept was well received. The near real-time filtering and immediate visual feedback was perceived as an important advantage over their existing tools, which usually take several minutes to calculate a result. However, it was mentioned that medical data is not complete: the initial parallel coordinates needed at least one value for each parameter. Missing data can be divided into three levels of impact on visualizations [33]: 1) perceivable impact, outliers will be visible immediately as a weight value of 734 kg, instead of 73.4 kg, will result in a line that deviates clearly from the other lines; 2) an invisible impact, especially in parallel coordinates, omitted/forgotten data can go unnoticed since no line is drawn; and a 3) propagating impact, when data is left out it can influence percentages and bias other data items. For these reasons, we adapted the parallel coordinates so that they can cope with missing values: a baseline is added underneath

the vertical axes to visualize missing values, as shown in Part A in Fig. 3. Three other possibilities were considered: 1) simply not visualize these patients, but many patients have at least one parameter that is never measured; 2) skip the specific missing parameter and thus have a broken line which makes it harder to see patterns in the parallel coordinates; and 3) use a dashed line which can easily be overlooked due to occlusion.

Furthermore, the demographic widgets of the first prototype were too detailed for GPs not working in any academic center and used too much screen estate, especially when the importance of secondary conditions and medication groups are considered. For example, diabetic patients who also suffer from a heart condition should receive medication from another group than 'regular' patients. This is why 'degree', 'religion', 'main GP' and 'income' are replaced with 'most frequent medication groups', 'most frequent medication', 'primary conditions' and 'secondary conditions' and moved to the table with the filtered patients.

C. No academic center affiliated general practice iteration

In the second iteration, the updated prototype was checked for perceived usefulness and usability at a general practice with three GPs not affiliated with any academic center. Before asking the GPs to elaborate on the design, the dashboard was shown and parallel coordinates were explained quickly.

The GPs shared their positive impression and stressed the benefits of using such a dashboard during their weekly medical team meeting. They especially appreciated the visualization of patients they saw in the previous week, which can easily be done by filtering patients on number of days since last visit.

D. Usability iteration

Since no ‘medical’ modifications were needed after the second iteration, combined with the fact that GPs have a very busy schedule and are thus hard to motivate for evaluation studies, the usability of the current prototype has formally been evaluated during a course on ‘information visualization’ at the University of Leuven, Belgium, with 12 students, a teaching assistant and the professor. Nevertheless, student are good candidates to identify usability issues [34], whereas perceived usefulness can only be tested with actual end users and is thus tested in the final iteration, as will be explained in more detail in Section V.

The usability evaluation lasted one hour and measured time-to-task, error-rate and perceived difficulty of tasks. In order to situate the results, the subjects were asked some initial questions to determine demographics and prior knowledge with visualizations. Secondly, they were asked to complete ten tasks, such as “How many patients who weight more than 60kg, take medication from the diuretics class?” and “What is the most common ‘secondary disease’ in the area bounded by Mechelen in the upper left corner and Leuven in the bottom right corner?”. For every task they had to register the start and completion time by adding an automatic time stamp in the electronic evaluation sheet and they had to indicate the difficulty of each task using a seven-point Likert scale ranging from 1) ‘very easy’ to 7) ‘very hard’. Next, they were asked general questions, such as “what does the red color mean” and “what did you like and did not like”? Seven out of ten tasks were completed in less than one minute and the task that took the most time (find the most common secondary disease in a particular area) took on average one minute and twenty-six seconds; this task was also rated the hardest task with a Likert score of 4.36, which is between ‘neither easy nor hard’ and ‘somewhat hard’. Furthermore, a System Usability Scale (SUS) [35] questionnaire was used to measure usability. The final score of 68 rates the usability of the application as ‘good’ [35], but clearly leaves room for improvements. Finally, the evaluation ended with an open discussion moment.

This evaluation exposed smaller bugs which could influence the usability, primarily due to performance, browser dependencies and different screen sizes. Several other suggestions were integrated into the next version of the design to improve the usability: 1) cleaner filtered patients table with the additional option to sort patients; 2) visual style update for the map in order not to distract GPs; and 3) consistent zoom and scroll controls.

E. Final design

The final design resulted, based on the complete set of evaluations, is shown in Fig. 3. Fig. 4 illustrates the overall process on a timeline displaying the different prototypes and most relevant changes made to each prototype.

IV. IMPLEMENTATION DETAILS

All prototypes are web-applications implemented in JavaScript using state-of-the-art open source toolkits, relying on the d3js framework [36] for visualizations. The parallel coordinates widget in the first prototype relied on the d3.parcoords toolkit⁶, whereas the other prototypes used our own implementation in order to visualize missing data and to add extra control options, such as the ability to highlight a single line (patient) when other filters are applied. The heatmap in the first three prototypes is built using the heatmap.js visualization library⁷, the heatmap in the final prototype is built using Google Maps integrated heatmap functionality.⁸ Thanks to the integration of the crossfilter library⁹ the large multi-variate dataset can be filtered with immediate visual feedback. Finally, as the tool is designed to be shown in full window mode, gridster.js¹⁰ takes care of the layout on every screen size and resolution. The source code of our tool can be found at: <https://github.com/AugmentHCI/ICHI2015>.

V. EVALUATION

The goal of the final evaluation was to assess the perceived usefulness and to assess the strengths, weaknesses, opportunities and threats of the proposed design.

A. Study subjects

Nine GPs (seven males and two females, ranging from 25 to 75 years old, and working in divergent parts of the country), were randomly recruited through mail or a phone call, in which we shortly explained the goal of this research and asked if they would want to participate in a short evaluation session. Participation was voluntary and not compensated. All subjects, but one, already use an EMR system.

B. Study protocol

Similar to O’Leary *et al.* [1], and as recommended by Taylor and Bogdan [37], an evaluation protocol was designed. GPs were not recorded since it could make them 1) feel uncomfortable [37]; 2) unwilling to discuss certain topics [1]; and 3) hold back information [1]. Throughout individual face-to-face interviews, the *concurrent Think-Aloud Protocol* [38] is used to let GPs explain what they think of when seeing the visualizations. In this way, it can be tested whether users understand the message the visualization tries to convey. Some specific questions are asked about the content, and the usability is quantified using a System Usability Scale (SUS) questionnaire [35]. Furthermore, we adapted the evaluation method from O’Leary *et al.* [1] in order to measure perceived usefulness, and thus also conducted a SWOT-analysis to assess Strengths, Weaknesses, Opportunities and Threats.

The idea was explained to each test subject and, if necessary, parallel coordinates were explained. Afterwards a demo was given to the GPs. In the remaining time the GPs could freely interact with the dashboard. Each evaluation

⁶<https://syntagmatic.github.io/parallel-coordinates/>

⁷<http://www.patrick-wied.at/static/heatmapjs/>

⁸<https://developers.google.com/maps/>

⁹<http://square.github.io/crossfilter/>

¹⁰<http://gridster.net/>

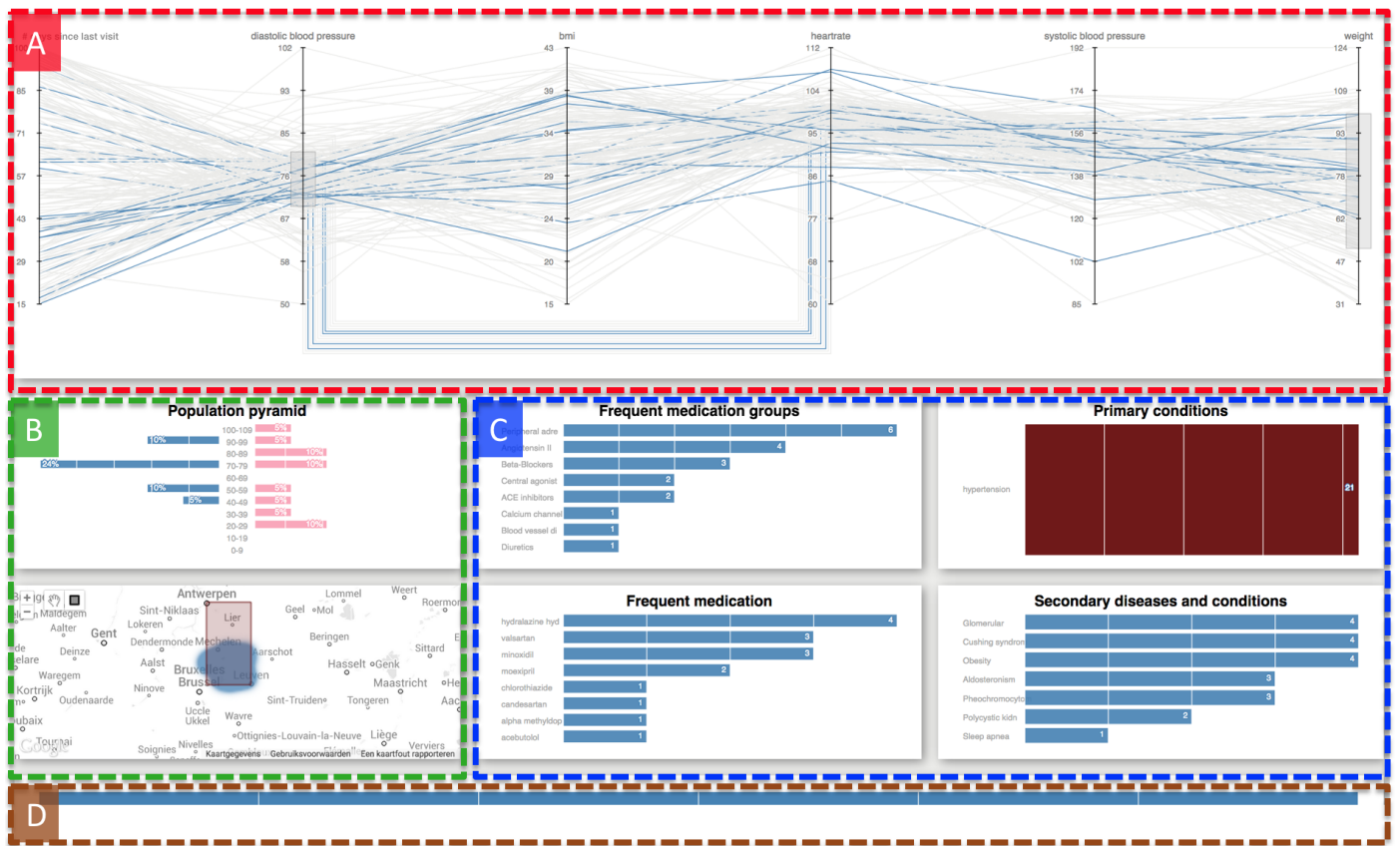


Fig. 3. Final design of the dashboard. In this example all patients with hypertension, a diastolic blood pressure between 70 - 80 mmHG and weight between 50 - 100 kg are selected. Furthermore, the resulting patients are also filtered by location: only patients with aforementioned values and that live in the area delimited by Antwerpen (Antwerp) and Leuven are selected. Part A shows the updated parallel coordinates widget, the horizontal lines indicate missing/never-measured values. The remaining demographic widgets are shown in part B. Part C contains the new medical filters: medication groups, medication, primary conditions, and secondary conditions. The red color indicates one category is selected in that particular widget, e.g. hypertension is selected as a primary condition. Finally, by showing a little part of the table, the user is hinted to scroll down in order to find the table that contains the selected patients.

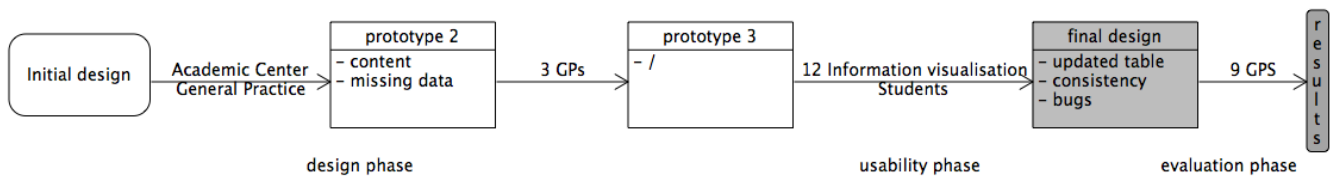


Fig. 4. Timeline that shows the prototypes. The initial design is evaluated within the Academic Center for General Practice in Leuven. The updated design is verified with three GPs not affiliated with any academic center. Since no changes were needed, the design is checked for usability with 12 information visualization students, a teaching assistant and the professor. Finally nine GPs evaluated the final design.

lasted at least 15 to 20 minutes which is right enough time for the explanation, the short demo and for the subjects to complete a SUS questionnaire and the questionnaire with Likert scale questions. The time limit was imposed by the busy schedule of the GPs. During the interview, the GPs were asked about the strong and weak points and what they see as possible opportunities or threats in order to complete the SWOT-analysis. The interview ended with a short open discussion.

C. Evaluation results

1) *Likert scale questions:* Usefulness questions were asked in a five-level Likert scale format ranging from 'totally disagree', 'disagree', 'neutral', 'agree' to 'totally agree'. As mentioned earlier, the questions are slightly adapted from [1], except for the last one which we added, and are listed below:

- Using a visual dashboard to improve patient follow-up is the right kind of tool.
- Interactive visualizations are a useful source of feedback.
- A visual dashboard can be effective in reducing

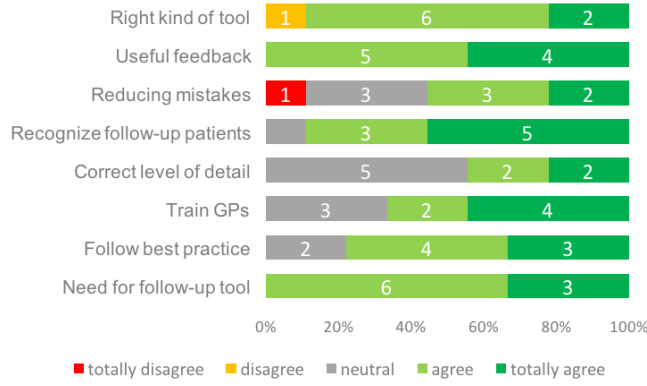


Fig. 5. Answers to the Likert scale questions. Green is positive, whereas orange and red mean a GP disagreed. Grey values indicate neutral answers. Questions based on [1].

mistakes in a clinical environment.

- A visual dashboard can be used to help GPs recognize patients that need follow-up.
- The details of the content shown on the visual dashboard are at the correct level.
- A visual dashboard is a good way to train GPs.
- A visual dashboard can be used to encourage GPs to follow best practice guidelines.
- There is a need for a tool for patient follow-up.

Fig. 5 provides a detailed overview of the distributions of the answers. Eight GPs out of nine think a visual dashboard is the right kind of tool to become aware about patients that need follow-up, whereas only one GP mentioned it was ‘a’ right tool, but not the only possible tool to gather insights into patient follow-up. All GPs agree that the design provides a useful source of feedback, that it can help to find patients who need a follow-up, and that there actually is a need for a patient follow-up tool. 6 out of 9 GPs think the dashboard can be used to train GPs and that it can be effective in reducing data registration errors. However, another GP mentioned it is probably more effective in correcting errors instead of preventing them. The dashboard can help to encourage following best practices according to 7 out of 9 GPs. Despite these results, only 4 out of 9 think the details of the contents shown are at the correct level.

2) *SWOT-analysis*: Table I outlines the strengths, weaknesses, opportunities and threats to a visual dashboard as identified by at least two test subjects. Five out of nine mentioned the ease with which they can select and filter patients. Four GPs saw the benefit of checking their patient data with evidence based guidelines, such as those made available by DuoDecim.¹¹ Furthermore, three GPs especially liked the visual overview and the ability to augment their work so lesser work is needed to perform a self-audit. Improved team communication thanks to the central overview, as well as the fact that no knowledge of logical operators is needed since the visual brushes take care of this transparently, are raised by two GPs. Although 4 out of 9 thought the map to be a weakness since it takes too much screen estate, which

TABLE I. SUMMARY OF THE SWOT-ANALYSIS. ONLY IDEAS AT LEAST TWO GPs MENTIONED ARE SHOWN.

Strengths	# GPs
- ease to select patients	5
- check with guidelines	4
- visual overview	3
- augment work	3
- improves team communication	2
- no pseudo code needed	2
Weaknesses	# GPs
- map uses too much screen estate	4
- not much structured data available	4
- not clear which content to show	3
- not needed often	2
Opportunities	# GPs
- ideal for research	5
- triggers self-reflection	4
- can be improved with patient collected data	3
Threats	# GPs
- averages can be dangerous	4
- too little time	3
- control system	2
- privacy	2

can be used for other medical information, 5 out of 9 do see the map as an opportunity for their own non-academic research. Another concern of 4 out of 9 GPs is that there is not enough structured data available to visualize, since GPs still register patient data as free text too often. Three GPs mention the fact that it is not always clear which content and parameters to shown on the dashboard. Finally, two out of nine think they would not need the dashboard often. Opportunities include: 1) useful for research, as mentioned by five out of nine; 2) improve self-reflection, as mentioned by four out of nine; and 3) add patient collected data, as mentioned by three out of nine. A big threat that surfaced during four interviews is the fact that only one value is displayed per parameter. At least three GPs feared that they would not have enough time to check their patient records. Unfortunately, since the government can ask for stats, 2 out of 9 GPs still think the tool can serve as some kind of control mechanism. Finally, although only an overview of the GP’s own records is visualized, privacy is mentioned by two GPs as a major threat.

3) *System Usability Scale questionnaire*: The GPs gave an average score of 74. The distribution of the different scores are displayed in Fig. 6 and discussed in Section VI.

VI. DISCUSSION

Initial feedback from end-users suggests that the dashboard can indeed be helpful in identifying patients in need of follow-up. Moreover, it can help discover patients who may otherwise be missed in the traditional EMR tools.

A. Likert scale questionnaire

As shown in Fig. 5, the Likert scale questions score rather positive: all GPs highlight the need for a tool to identify patients in need of follow-up visits. Furthermore, all GPs like the feedback that the dashboard provides and think it can help them recognize patients who need follow-up. Only one GP thinks interactive visualizations are not necessarily the only

¹¹<http://www.duodecim.fi/web/english/home>

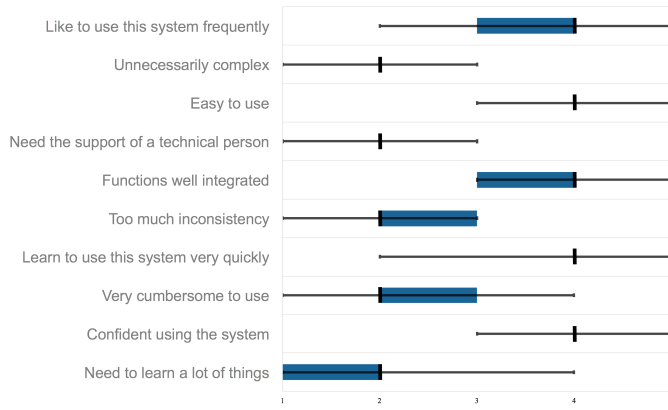


Fig. 6. The distribution of the SUS scores on the different questions. A high score means the user completely agrees with the statement. The SUS questions are ordered to create alternating scores in ideal score and range from 1 ‘totally disagree’ to 5 ‘totally agree’. The questions are shortened in this figure for readability. For a more complete description and interpretations of the SUS questions the reader is referred to Bangor *et al.* [35].

right kind of tool, and that other tools could assist, such as data mining techniques or “clever algorithms” [5]. Nevertheless, it can be concluded that a visual dashboard is still ‘a’ right tool for patient follow-up insights. The neutral to positive answers to the level of detail shown can be explained by the fact that only the most common parameters were shown in the proof-of-concept. This concern was also repeated as a weakness in the SWOT-analysis. A real-life dashboard should allow GPs to choose the visualized parameters according to what they want to research. One GP totally disagreed that an overview tool can help to reduce medical errors in a medical setting; “*the dashboard can definitely help to find and correct errors*”, yet the actual prevention should be done in an earlier stage. This is true and it can be concluded that the design can help to expose (and correct) medical mistakes, but cannot necessary help to prevent them, although it could raise awareness to enter data in a more rigorous way.

B. SWOT-analysis

According to the GPs in our study, the most important strength is the ease with which patients can be filtered. The ability to easily and dynamically select - or brush - is in fact an important part of interactive visualizations [39]. The near real-time brushing is considered a major advantage over the existing tools. An often repeated remark is that the dashboard can easily be used to check one’s performance on evidence based guidelines, to see how patient records deviate from standards. Finally, as dashboards can be used to improve communication between collaborators [40], it was mentioned by GPs that our dashboard can improve team communication in meetings, which was also mentioned by the three GPs in the intermediate iteration; an interesting set-up would be to visualize patients the GPs processed since their previous team meeting on a shared big screen.

The geographical map is considered the ‘weakest’ element by four out of nine test subjects. Geographical brushing can be useful, yet the heatmap is primarily useful for academic research. Another concern of the GPs is that there is not enough structured data available, since GPs still register

patient data as free text too often. Furthermore, handling unstructured medical data is still an open research question [5]. Nevertheless, GPs are slowly getting used to registering data in a structured way. A wanted side-effect of the dashboard is to provide an additional motivation to start registering in a structured manner.

The proof-of-concept is seen as having potential in academic research, both the Academic Center for General Practice of Leuven and the more research-minded GPs saw the benefits of using the dashboard to analyze patient files and discover possible patterns. It could - and also should - trigger self-reflection, yet this needs to be validated in a real-life test setting. When data would be augmented with patient collected data, for example from Apple’s HealthKit, the design has even more potential. Using interactive visualizations, GPs will be able to analyze more data in less time. Nevertheless, before making strong claims, this opportunity should be further explored in future research.

The biggest threat is the fact that only one value per parameter can be visualized. Especially in a medical context, showing only average values can be dangerous. For instance, a previous extremely high diastolic blood pressure value of 120 mmHg and an extremely low value of 40 mmHg still give a normal average of 80 mmHg. An easy solution would be to only show the last value. Unfortunately, this way a lot of information is lost, since it probably means something if blood pressure values differ this much. Another threat as raised by some GPs is the small amount of free time GPs have for such reflection moments, as we also experienced when contacting participants: GP workload is very high [41]. Although the dashboard enables GPs to perform an audit themselves without someone controlling their work, some GPs still fear the government may start requiring these results in order to receive funding. Finally, similar to Dasgupta *et al.* [8], privacy is mentioned by our test subjects as an attention point. However, the dashboard only visualizes data the GP already has access to so we consider this a non-important in this particular case.

C. System Usability Scale questionnaire

The average SUS score of 74 rates the application between ‘good’ and ‘excellent’ [35]. Nevertheless, since this average is only based on nine participants, no final conclusions can be derived from this score. We therefore looked at the individual questions, as shown in Fig. 6, and found some interesting results. All questions score good with a median of 2 where a low score is expected and of 4 where a high score is expected. The first question where GPs are asked if they would like to use this system frequently scores positives, yet the weight of the answer lies between 3 and 4. However, this is likely due to the word ‘frequently’ since GPs indicated they would primarily use our dashboard during team meetings or during reflection moments. The learning curve of parallel coordinates is clearly not considered a big issue by our test subjects since questions 7 and 10, which address the learning curve, score as desired.

D. Limitations

1) *Sample size*: Our design was evaluated with, in total, 15 GPs of which nine in the final iteration. This is a relatively

small sample size, yet it enabled us to sufficiently show the potential of the proof-of-concept. We should now integrate it into a real EMR which will allow us to gather more detailed user analytics. Furthermore, it is also a relative standard sample size in this specific domain. O’Leary *et al.* [1] have in total 19 participants of which nine nurses, and only six GPs. Hirsch *et al.* [42] used twelve physicians in their formative evaluation and Zhang *et al.* [23] only use six physicians and two health informatics professionals. Finally, in usability research, five users are generally enough to find 85% of usability problems [43].

2) *Semi realistic data*: The proof-of-concepts are based on realistic medical data, but only pre-selected random data is shown in order to make the data anonymous. This prevents the detection of actual patterns, and might negatively influence the perception on the proof-of-concept, since absurd data or relations might be visualized. Nevertheless, using random data still allowed us to assess the perceived usefulness. In future work, it should be researched if visualizing the GPs own patient records improves the perception and if they are able to find patterns in the data.

3) *Requirements gathering*: No formal requirements gathering phase was organized. The initial design is based on knowledge gained during informal talks with stakeholders, such as GPs, and on theoretical concepts as discussed in Section III-A. However, the user centered, rapid prototyping methodology with real GPs ensures the feedback of the user is integrated. Yet, some delays, like the use of average values, could have been avoided using a more formal initial requirement gathering phase. A framework such as the nine-stage design study methodology framework from Sedlmair, Meyer, and Munzner [44] would have prevented these delays.

4) *Evaluation setting*: For this evaluation setting we followed similar work published in high quality conferences and journals such as the work of Hirsch *et al.* [42], O’Leary *et al.* [1], Ratwani and Fong [16], and Zhang *et al.* [23]. Since this work is explicitly stated as an exploratory study, the proof-of-concept dashboard is not evaluated ‘in the wild’. The design is only evaluated through short face-to-face interviews in Belgium and the conclusions are based on these interviews. Nevertheless, with this evaluation setting we were able to show the need for a follow-up tool. Furthermore, our results demonstrate the strengths, weaknesses, opportunities, and threats an interactive visual dashboard for patient follow-up has to offer.

VII. CONCLUSION AND FUTURE WORK

Targeted follow-up meetings are important in general practice and are missed too often. Our proof-of-concept dashboard provides GPs with a powerful, yet easy to use method to analyze more easily the numerous patient files. It reduces the burden of analyzing patient records and, consequently, is likely to encourage greater data exploration and improve the discovery of meaningful trends. Moreover, the dashboard is not only useful in our specific niche of patient follow-up in general practice, the benefits for medical researchers were often stressed by our test subjects.

The dashboard is designed following a user centered, rapid prototyping methodology and evaluated using an evaluation framework as followed by O’Leary *et al.* [1]. Results indicate that there is a need for a follow-up tool and a visual dashboard, compared to the more traditional tools, is the preferred option by eight out of nine interviewed GPs. Our proof-of-concept is able to show useful feedback and can help GPs recognize patients in need of a follow-up. The major strengths are the ease with which GPs can query patients and check how they are performing compared to evidence based guidelines. Furthermore, the biggest opportunities are in research, yet GPs also indicated that our visualization can trigger self-reflection on follow-up capabilities. Finally, an average SUS score of 74 indicates that the usability of our implementation can be considered as ‘good’.

In future work, we would like to investigate the possibility of connecting the design to electronic medical records in order to directly visualize GPs’ patient records and find realistic patterns. We also hope to make the dashboard available in general practices, so that it can be used ‘in the wild’, during medical team meetings in order to evaluate the impact on collaboration and to see if self-reflection is triggered.

VIII. ACKNOWLEDGMENTS

This research was funded under research grand IWT 120896. The authors would like to thank Corilus for access to their products and the Academic Center for General Practice in Leuven, the students, and all GPs for extensive feedback and input. Furthermore, the authors would also like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

REFERENCES

- [1] P. O’Leary, N. Carroll, and I. Richardson, “The Practitioner’s Perspective on Clinical Pathway Support Systems,” in *2014 IEEE International Conference on Healthcare Informatics*, pp. 194–201.
- [2] S. Bird, “A GP’s duty to follow up test results,” *Australian family physician*, vol. 32, no. 1-2, pp. 45–46, 2003.
- [3] I. Ganguli, J. H. Wasfy, and T. G. Ferris, “What Is the Right Number of Clinic Appointments?” *Journal of the American Medical Association*.
- [4] D. Engelbart, “Augmenting Human Intellect: A Conceptual Framework,” Tech. Rep., 1962. [Online]. Available: <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0289565>
- [5] J. J. Caban and D. Gotz, “Visual analytics in healthcare – opportunities and research challenges,” *Journal of the American Medical Informatics Association*, vol. 22, no. 2, pp. 260–262, 2015.
- [6] A. Rind, “Interactive Information Visualization to Explore and Query Electronic Health Records,” *Foundations and Trends in Human-Computer Interaction*, no. 3, pp. 207–298.
- [7] V. L. West, D. Borland, and W. E. Hammond, “Innovative information visualization of electronic health record data: a systematic review,” *Journal of the American Medical Informatics Association*, no. 2, pp. 330–339.
- [8] A. Dasgupta, E. Maguire, A. Abdul-rahman, and M. Chen, “Opportunities and Challenges for Privacy-Preserving Visualization of Electronic Health Record Data,” in *Proc. of IEEE VIS 2014 Workshop on Visualization of Electronic Health Records*, Paris, 2014.
- [9] S. Malik, F. Du, M. Monroe, E. Onukwugha, C. Plaisant, and B. Shneiderman, “An Evaluation of Visual Analytics Approaches to Comparing Cohorts of Event Sequences,” in *Proc. of IEEE VIS 2014 Workshop on Visualization of Electronic Health Records*, Paris, 2014, pp. 1–6.

- [10] —, “Cohort Comparison of Event Sequences with Balanced Integration of Visual Analytics and Statistics,” in *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15*, pp. 38–49.
- [11] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman, “LifeLines: visualizing personal histories,” in *Proceedings of the (SIGCHI) conference on Human factors in computing systems: common ground*, pp. 221–227.
- [12] T. D. Wang, C. Plaisant, B. Shneiderman, N. Spring, D. Roseman, G. Marchand, V. Mukherjee, and M. Smith, “Temporal summaries: Supporting temporal categorical searching, aggregation and comparison,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1049–1056, 2009.
- [13] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman, “Temporal event sequence simplification,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2227–2236, 2013.
- [14] D. Klimov, Y. Shahar, and M. Taieb-Maimon, “Intelligent visualization and exploration of time-oriented data of multiple patients,” *Artificial Intelligence in Medicine*, no. 1, pp. 11–31.
- [15] D. Gotz, J. Sun, N. Cao, and S. Ebadollahi, “Visual cluster analysis in support of clinical decision intelligence,” *AMIA Annual Symposium proceedings*, pp. 481–90.
- [16] R. M. Ratwani and A. Fong, “‘Connecting the dots’: leveraging visual analytics to make sense of patient safety event reports,” *Journal of the American Medical Informatics Association*, pp. 312–317.
- [17] K. Wongsuphasawat and D. Gotz, “Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2659–2668, 2012.
- [18] P. Riehmann, M. Hanfler, and B. Froehlich, “Interactive sankey diagrams,” in *Proceedings - IEEE Symposium on Information Visualization, INFO VIS*, 2005, pp. 233–240.
- [19] K. Wongsuphasawat and D. Gotz, “Outflow: Visualizing patient flow by symptoms and outcome,” *Visual Analytics in Healthcare, IEEE VisWeek Workshop on*, no. March, pp. 25–28, 2011.
- [20] a. Inselberg and B. Dimsdale, “Parallel coordinates: a tool for visualizing multi-dimensional geometry,” 1990.
- [21] B. Jodeiry, M. Heidarzadeh, K. Mirmia, F. Akrami, S. Heidarabadi, and A. Ebadi, “Innovation of high-risk infants follow-up surveillance system in Iran,” *International Journal of Preventive Medicine*, vol. 6, no. 1, p. 35, 2015.
- [22] D. Dowding, R. Randell, P. Gardner, G. Fitzpatrick, P. Dykes, J. Favela, S. Hamer, Z. Whitewood-Moores, N. Hardiker, E. Borycki, and L. Currie, “Dashboards for improving patient care: Review of the literature,” *International Journal of Medical Informatics*, vol. 84, no. 2, pp. 87–100, Feb. 2015.
- [23] Z. Zhang, B. Wang, F. Ahmed, I. V. Ramakrishnan, R. Zhao, A. Viccellio, and K. Mueller, “The five Ws for information visualization with application to healthcare informatics,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 11, pp. 1895–1910, 2013.
- [24] R. De Croon, J. Klerkx, and E. Duval, “Designing a Useful and Usable Mobile EMR Application through a Participatory Design Methodology: A Case Study,” in *2014 IEEE International Conference on Healthcare Informatics*, Verona, pp. 176–185.
- [25] B. Shneiderman, “The eyes have it: a task by data type taxonomy for information visualizations,” in *Proceedings 1996 IEEE Symposium on Visual Languages*. Boulder: IEEE, pp. 336 – 343.
- [26] —, “Dynamic queries for visual information seeking,” *IEEE Software*, vol. 11, no. 6, pp. 70–77, 1994.
- [27] S. Stevens, “On the theory of scales of measurement,” *Science (New York, N.Y.)*, vol. 103, no. 2684, pp. 677–680, 1946.
- [28] O. von dem Knesebeck, H. Bickel, A. Fuchs, J. Gensichen, S. Höfels, S. G. Riedel-Heller, H.-H. König, K. Mergenthal, G. Schön, K. Wegscheider, S. Weyerer, B. Wiese, M. Scherer, H. van den Bussche, and I. Schäfer, “Social inequalities in patient-reported outcomes among older multimorbid patients – results of the MultiCare cohort study,” *International Journal for Equity in Health*, no. 1, pp. 1–11.
- [29] J. Heinrich and D. Weiskopf, “State of the Art of Parallel Coordinates,” in *Eurographics*, 2013, pp. 95–116.
- [30] S. P. van Noort, M. Muehlen, H. Rebelo de Andrade, C. Koppeschaar, J. M. Lima Lourenco, and M. G. M. Gomes, “Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe,” *Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, vol. 12, no. 7, pp. E5–6, Jul. 2007.
- [31] E. R. Tufte, *The visual display of quantitative information*, second ed. ed. Cheshire, Connecticut: Graphics Press.
- [32] V. Dudas, T. Bookwalter, K. M. Kerr, and S. Z. Pantilat, “The impact of follow-up telephone calls to patients after hospitalization,” *The American Journal of Medicine*, no. 9, pp. 26–30, Apr.
- [33] C. Eaton, C. Plaisant, and T. Drizd, “Visualizing missing data: Graph interpretation user study,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3585 LNCS, 2005, pp. 861–872.
- [34] S. Kriglstein and M. Pohl, “Choosing the Right Sample? Experiences of Selecting Participants for Visualization Evaluation,” in *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3)*, W. Aigner, P. Rosenthal, and C. Scheidegger, Eds. The Eurographics Association, 2015.
- [35] A. Bangor, P. Kortum, and J. Miller, “Determining what individual SUS scores mean: Adding an adjective rating scale,” *Journal of usability studies*, no. 3, pp. 114–123.
- [36] M. Bostock, V. Ogievetsky, and J. Heer, “D3 data-driven documents,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2301–2309, 2011.
- [37] S. J. Taylor and R. Bogdan, *Introduction to Qualitative Research Methods*. John Wiley and Sons.
- [38] C. Lewis, *Using the “thinking Aloud” Method in Cognitive Interface Design*, (ibm resea ed., ser. Research report. Yorktown Heights, NY: IBM T.J. Watson Research Center.
- [39] J. Heer, B. Shneiderman, and C. Park, “A taxonomy of tools that support the fluent and flexible use of visualizations,” *Interactive Dynamics for Visual Analysis*, pp. 1–26.
- [40] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale, “Empirical studies in information visualization: Seven scenarios,” pp. 1520–1536, 2012.
- [41] “British Medical Association National survey of GPs The future of General Practice 2015 Full Report,” Tech. Rep., 2015. [Online]. Available: <http://bma.org.uk/working-for-change/negotiating-for-the-profession/bma-general-practitioners-committee/surveys/future-of-general-practice>
- [42] J. S. Hirsch, J. S. Tanenbaum, S. Lipsky Gorman, C. Liu, E. Schmitz, D. Hashorva, A. Ervits, D. Vawdrey, M. Sturm, and N. Elhadad, “HARVEST, a longitudinal patient record summarizer,” *Journal of the American Medical Informatics Association*, pp. 263–274.
- [43] J. Nielsen, “Why You Only Need to Test with 5 Users,” 2000. [Online]. Available: <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- [44] M. Sedlmair, M. Meyer, and T. Munzner, “Design Study Methodology: Reflections from the Trenches and the Stacks,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2431–2440, Dec. 2012.